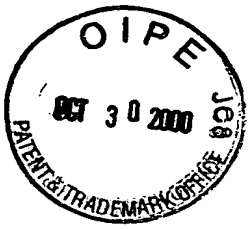


UNITED STATES PATENT AND TRADEMARK OFFICE

I, Susan POTTS BA ACIS,

Director of RWS Group plc, of Europa House, Marsham Way, Gerrards Cross, Buckinghamshire, England declare;

1. That I am a citizen of the United Kingdom of Great Britain and Northern Ireland.

2. That the translator responsible for the attached translation is well acquainted with the German and English languages.

3. That the attached is, to the best of RWS Group plc knowledge and belief, a true translation into the English language of the accompanying copy of the specification filed with the application for a patent in Germany on 4 May 1998 under the number 198 19 829.9 and the official certificate attached hereto.

4. That I believe that all statements made herein of my own knowledge are true and that all statements made on information and belief are true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the patent application in the United States of America or any patent issuing thereon.

For and on behalf of RWS Group plc

The 23rd day of October 2000

FEDERAL REPUBLIC OF GERMANY

[Eagle crest]

Priority Certificate

for the filing of a Patent Application

File Reference: 198 19 829.9

Filing date: 4 May 1998

Applicant/Proprietor: Bayer Aktiengesellschaft, Leverkusen/DE

Title: Nucleic acids which encode insect acetylcholine receptor subunits

IPC: C 07 K, A 61 K, C 12 N

The attached documents are a correct and accurate reproduction of the original submission for this Application.

Munich, 26 September 2000

German Patent and Trademark Office

The President

pp

[Seal of the German Patent and Trademark Office]

[signature]

Dzierzon

Nucleic acids which encode insect acetylcholine receptor subunits

The invention relates, in particular, to nucleic acids which encode insect acetylcholine receptor subunits.

Nicotinic acetylcholine receptors are ligand-regulated ion channels which are of importance in neurotransmission in the animal kingdom. The binding of acetylcholine or other agonists to the receptor induces a transient opening of the channel and allows cations to flow through. It is assumed that a receptor consists of five subunits which are grouped around a pore. Each of these subunits is a protein which consists of an extracellular N-terminal moiety followed by three transmembrane regions, an intracellular moiety, a fourth transmembrane region and a short extracellular C-terminal moiety (Changeux et al. 1992).

Acetylcholine receptors are especially well investigated in vertebrates. In this context, three groups can be distinguished on the basis of their anatomical location and their functional properties (conducting properties of the channel, desensitization, and sensitivity towards agonists and antagonists and also towards toxins such as α–bungarotoxin). The classification correlates with the molecular composition of the receptors. There are heterooligomeric receptors having the subunit composition $\alpha_2\beta\gamma\delta$, which are found in muscle (Noda et al. 1982, Claudio et al. 1983, Devillers-Thiery et al. 1983, Noda et al. 1983a, b), heterooligomeric receptors which contain subunits from the α2 - α6 and β2 - β4 groups and which are found in the nervous system (Wada et al. 1988, Schoepfer et al. 1990, Cockcroft et al. 1991, Heinemann et al. 1997), and also homooligomeric receptors which contain subunits from the α7 - α9 group and which are likewise found in the nervous system (Lindstrom et al. 1997, Elgoyhen et al. 1997). This classification is also supported by an examination of the relatedness of the gene sequences of the different subunits. Typically, the sequences of functionally homologous subunits from different species are more similar to each other than are sequences of subunits which are from different groups but from the same species. Thus, the rat muscle α subunit, for example, exhibits 78% amino acid identity and 84% amino acid similarity with that of the electric ray Torpedo californica but only 48% identity and 59% similarity with the rat α2 subunit (heterooligomeric, neuronal) and 36% identity and 45% similarity with the rat α7 subunit (homooligomeric, neuronal). Furthermore, the gene sequences of all the

known acetylcholine receptor subunits are to a certain extent similar not only to each other but also to those of some other ligand-regulated ion channels (e.g. the serotonin receptors of the $5HT_3$ type, the GABA-regulated chloride channels and the glycine-regulated chloride channels). It is therefore assumed that all these receptors are descended from one common precursor and they are classified into one supergene family (Ortells et al. 1995).

In insects, acetylcholine is the most important excitatory neurotransmitter of the central nervous system. Accordingly, acetylcholine receptors can be detected electrophysiologically in preparations of insect central nervous system ganglia. The receptors are detected both in postsynaptic and presynaptic nerve endings and in the cell bodies of interneurones, motor neurones and modulatory neurones (Breer et al. 1987, Buckingham et al. 1997). Some of the receptors are inhibited by α-bungarotoxin while others are insensitive (Schloß et al. 1988). In addition, the acetylcholine receptors are the molecular point of attack for important natural (e.g. nicotine) and synthetic insecticides (e.g. chloronicotinyls).

The gene sequences of a number of insect nicotinic acetylcholine receptors are already known. Thus, the sequences of five different subunits have been described in Drosophila melanogaster (Bossy et al. 1988, Hermanns-Borgmeyer et al. 1986, Sawruk et al. 1990a, 1990b, Schulz et al. Unpublished, EMBL accession number Y15593), while five have likewise been described in Locusta migratoria (Stetzer et al. unpublished, EMBL accession numbers AJ000390 - AJ000393), one has been described in Schistocerca gregaria (Marshall et al. 1990), two have been described in Myzus persicae (Sgard et al. unpublished, EMBL accession number X81887 and X81888), and one has been described in Manduca sexta (Eastham et al. 1997). Furthermore, a number of partial gene sequences from Drosophila melanogaster have been characterized as so-called expressed sequence tags (Genbank accession numbers AA540687, AA698155, AA697710, AA697326). The fact that individual sequences are very similar to those from other insects suggests that these subunits are functional homologues.

It is of great practical importance to make available new insect acetylcholine receptor subunits, for example for the purpose of searching for novel insecticides, with those subunits which differ from the known subunits to a greater extent than is the case between functional homologues being particularly of interest.

The present invention is consequently based, in particular, on the object of making available nucleic acids which encode novel insect acetylcholine receptor subunits.

This object is achieved by the provision of nucleic acids which comprise a sequence selected from

(a) the sequences according to SEQ ID NO: 1, SEQ ID NO: 3 or SEQ ID NO: 5,

(b) part sequences of the sequences defined in (a) which are least 14 base pairs in length,

(c) sequences which hybridize to the sequences defined in (a) in 2 x SSC at 60°C, preferably in 0.5 x SSC at 60°C, particularly preferably in 0.2 x SSC at 60°C (Sambrook et al. 1989),

(d) sequences which exhibit at least 70% identity with the sequences defined in (a), between position 1295 and position 2195 in the case of SEQ ID NO: 1, or between position 432 and position 1318 in the case of SEQ ID NO: 3, or between position 154 and position 1123 in the case of SEQ ID NO: 5,

(e) sequences which are complementary to the sequences defined in (a), and

(f) sequences which, because of the degeneracy of the genetic code, encode the same amino acid sequences as the sequences defined in (a) to (d).

The degree of identity of the nucleic acid sequences is preferably determined using the GAP program from the GCG program package, Version 9.1 with standard settings.

The present invention is based on the surprising finding that insects possess genes which encode subunits of, in particular, homooligomeric acetylcholine receptors.

The invention furthermore relates to vectors which contain at least one of the novel nucleic acids. All the plasmids, phasmids, cosmids, YACs or artificial chromosomes which are used in molecular biological laboratories can be used as vectors. These vectors can be linked to the usual regulatory sequences for the purpose of expressing the novel nucleic acids. The choice of such regulatory sequences depends on whether prokaryotic or eukaryotic cells, or cell-free systems, are used for the expression. The SV40, adenovirus or cytomegalovirus early or late promoter, the lac system, the trp system, the main operator and promoter regions of phage lambda, the control regions of the fd coat protein, the 3-phosphoglycerate kinase promoter, the acid phosphatase promoter and the yeast α-mating factor promoter are examples of expression control sequences which are particularly preferred.

In order to be expressed, the nucleic acids according to the invention can be introduced into suitable host cells. Both prokaryotic cells, preferably E.coli, and eukaryotic cells, preferably mammalian or insect cells, are suitable for use as host cells. Other examples of suitable unicellular host cells are: Pseudomonas, Bacillus, Streptomyces, yeasts, HEK-293, Schneider S2, CHO, COS1 and COS7 cells, plant cells in cell culture and also amphibian cells, in particular oocytes.

The present invention also relates to polypeptides which are encoded by the nucleic acids according to the invention and also the acetylcholine receptors, preferably homooligomeric acetylcholine receptors, which are synthesized from them.

In order to prepare the polypeptides which are encoded by the nucleic acids according to the invention, host cells which contain at least one of the nucleic acids according to the invention can be cultured under suitable conditions. After that, the desired polypeptides can be isolated from the cells or the culture medium in a customary manner.

The invention furthermore relates to antibodies which bind specifically to the abovementioned polypeptides or receptors. These antibodies are prepared in the customary manner. For example, such antibodies can be produced by injecting a substantially immunocompetent host with a quantity of an acetylcholine receptor polypeptide, or a fragment thereof, according to the invention which is effective for producing antibodies, and subsequently isolating these antibodies. Furthermore, an immortalized cell line which produces monoclonal antibodies can be obtained in a

manner known per se. Where appropriate, the antibodies can be labelled with a detection reagent. Preferred examples of such a detection reagent are enzymes, radioactively labelled elements, fluorescent chemicals or biotin. Instead of the complete antibody, use can also be made of fragments which possess the desired specific binding properties.

The nucleic acids according to the invention can be used, in particular, for producing transgenic invertebrates. These latter can be employed in test systems which are based on an expression of the receptors according to the invention, or variants thereof, which differs from that of the wild type. In addition, this includes all transgenic invertebrates in which a change in the expression of the receptors according to the invention, or their variants, occurs as the result of modifying other genes or gene control sequences (promoters).

The transgenic invertebrates are produced, for example, in Drosophila melanogaster by means of P element-mediated gene transfer (Hay et al., 1997) or in Caenorhabditis elegans by means of transposon-mediated gene transfer (e.g. using Tc1, Plasterk, 1996).

The invention also consequently relates to transgenic invertebrates which contain at least one of the nucleic acid sequences according to the invention, preferably to transgenic invertebrates of the species Drosophila melanogaster or Caenorhabditis elegans, and to their transgenic progeny. Preferably, the transgenic invertebrates contain the receptors according to the invention in a form which differs from that of the wild type.

The nucleic acids according to the invention can be prepared in the customary manner. For example, the nucleic acid molecules can be synthesized entirely chemically. In addition, only short segments of the sequences according to the invention can be synthesized chemically and these oligonucleotides can be labelled radioactively or with a fluorescent dye. The labelled oligonucleotides can be used to screen cDNA libraries prepared from insect mRNA. Clones which hybridize to the labelled oligonucleotides are selected for isolating the relevant DNA. After the isolated DNA has been characterized, the nucleic acids according to the invention are readily obtained.

The nucleic acids according to the invention can also be prepared by means of PCR methods using chemically synthesized oligonucleotides.

The nucleic acids according to the invention can be used for isolating and characterizing the regulatory regions which occur naturally adjacent to the coding region. Consequently, the present invention also relates to these regulatory regions.

The nucleic acids according to the invention can be used to identify novel active compounds for plant protection, such as compounds which, as modulators, in particular as agonists or antagonists, alter the conducting properties of the acetylcholine receptors according to the invention. For this, a recombinant DNA molecule, which encompasses at least one nucleic acid according to the invention, is introduced into a suitable host cell. The host cell is cultured, in the presence of a compound or a sample which comprises a multiplicity of compounds, under conditions which permit expression of the receptors according to the invention. A change in the receptor properties can be detected, as described below in Example 2. Using this approach, it is possible to discover insecticidal substances.

The nucleic acids according to the invention also make it possible to discover compounds which bind to the receptors according to the invention. These compounds can likewise be used as insecticides on plants. For example, host cells which contain the nucleic acid sequences according to the invention and express the corresponding receptors or polypeptides, or the gene products themselves, are brought into contact with a compound or a mixture of compounds under conditions which permit the interaction of at least one compound with the host cells, receptors or the individual polypeptides.

Host cells or transgenic invertebrates with contain the nucleic acids according to the invention can also be used to discover substances which alter the expression of the receptors.

The above-described nucleic acids, vectors and regulatory regions according to the invention can additionally be used for discovering genes which encode polypeptides which are involved in the synthesis, in insects, of functionally similar acetylcholine receptors. According to the present invention, functionally similar receptors are understood as being receptors which encompass polypeptides which, while differing in their amino acid sequences from the polypeptides described in this present publication, essentially possess the same functions.

Comments on the sequence listing and the figures:

SEQ ID NO: 1 shows the nucleotide sequence of the isolated Da7 cDNA, beginning with position 1 and ending with position 2886. SEQ ID NO: 1 and SEQ ID NO: 2 also show the amino acid sequences of the protein deduced from the Da7 cDNA sequence.

SEQ ID NO: 3 shows the nucleotide sequence of the isolated Hva7-1 cDNA, beginning with position 1 and ending with position 3700. SEQ ID NO: 3 and SEQ ID NO: 4 also show the amino acid sequences of the protein deduced from the Hva7-1 cDNA sequence.

SEQ ID NO: 5 shows the nucleotide sequence of the isolated Hva7-2 cDNA, beginning with position 1 and ending with position 3109. SEQ ID NO: 5 and SEQ ID NO: 6 also show the amino acid sequences of the protein deduced from the Hva7-2 cDNA sequence.

Figure 1 shows the increase in intracellular calcium which occurs in cells which have been recombinantly modified as described in Example 2 following the addition of nicotine. Cells were loaded with Fura-2-acetoxymethyl ester (5 - 10 μM in serum-free minimal essential medium containing 1% bovine serum albumin and 5 mM calcium chloride), washed with Tyrode solution buffered with N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulphonic acid) (5 mM HEPES) and alternately illuminated, under a fluorescence microscope (Nikon Diaphot) with light of 340 nm and 380 nm wavelength. A measurement point corresponds to a pair of video images at the two wavelengths (exposure time per image, 100 ms). The time interval between two measurement points is 3 s. After 8 images had been taken (measurement point 4.0), nicotine was added to a final concentration of 500 μM and the measurement series was continued. The fluorescence intensity of the cells when illuminated with light of 380 nm wavelength was divided by the corresponding intensity at 340 nm, thereby giving the ratio.

Examples:

Example 1

Isolating the described polynucleotide sequences

Polynucleotides were manipulated using standard methods of recombinant DNA technology (Sambrook, et al., 1989). The bioinformatic processing of nucleotide and protein sequences was carried out using the GCG program package Version 9.1 (GCG Genetics Computer Group, Inc., Madison Wisconsin, USA).

Partial polynucleotide sequences

Sequence comparisons ("Clustalw") were used to identify regions, from which degenerate oligonucleotides were deduced by backtranslating the codons, of protein sequences from genes whose ability to form homooligomeric acetylcholine receptors was known. In all, 5 such oligonucleotide pairs were selected for the polymerase chain reaction (PCR). Only one combination (see below) gave a product both from Heliothis cDNA and from Drosophila cDNA.

RNA was isolated from whole Heliothis virescens embryos (shortly before hatching) using Trizol reagent (Gibco BRL, in accordance with the manufacturer's instructions). The same procedure was adopted with Drosophila embryos (24 h at 25°C). 10 μg of these RNAs were employed in a first cDNA strand synthesis (Superscript Preamplification System for first cDNA strand synthesis, Gibco BRL, in accordance with the manufacturer's instructions, reaction temperature 45°C).

Subsequently, 1/100 of the abovementioned first-strand cDNA was in each case employed in a polymerase chain reaction (PCR) using the oligonucleotides alpha7-1s: (5'-GAYGTIGAYGARAARAAYCA-3') and alpha7-2a: (5'-CYYTCRTCIGCRCTRTTRTA-3') (recombinant Taq DNA polymerase, Gibco BRL). The PCR parameters were as follows: Hva7-1 and Hva7-2: 94°C, 2 min; 35 times (94°C, 45 s; 50°C, 30 s; 72°C, 60 s) and also Da7: 96°C, 2 min; 35 times (96°C, 45 s; 50°C, 30 s; 72°C, 60 s). In each case, this resulted in a dectable band of approx. 0.2 kb in an agarose gel (1%), both in the case of Drosophila cDNA and in the case of Heliothis cDNA. After the DNA fragments had been subcloned by means

of SrfScript (Stratagene), and their sequences had been determined, it turned out that two different DNA fragments had been amplified from Heliothis cDNA; these were 228-11 = Hva7-1 (partial, containing 165 bp) and 228-8 = Hva7-2 (partial, containing 171 bp). Only one DNA fragment was isolated from Drosophila cDNA; this was 248-5 = Da7 (partial, containing 150 bp).

Isolating poly A-containing RNA from Heliothis virescens tissue and constructing the cDNA libraries

The RNA for cDNA library I was isolated from whole Heliothis virescens embryos (shortly before hatching) using Trizol reagent (Gibco BRL, in accordance with the manufacturer's instructions). The RNA for cDNA library II was isolated from whole head ganglia from 500 Heliothis virescens larvae (stages 4-5) usings Trizol reagent (Gibco BRL, in accordance with the manufacturer's instructions). The poly A-containing RNAs were then isolated from these RNAs by purifying with Dyna Beads 280 (Dynal). 5 μg of these poly A-containing RNAs were subsequently employed in constructing cDNA libraries I and II using the λ-ZAPExpress vector (cDNA Synthesis Kit, ZAP-cDNA Synthesis Kit and ZAP-cDNA Gigapack III Gold Cloning Kit, all from Stratagene). In a departure from the manufacturer's instructions, Superscript Reverse Transcriptase (Gibco BRL) was used for synthesizing the cDNA at a synthesis temperature of 45°C. In addition, radioactively labelled deoxynucleoside triphosphates were not added. Furthermore, the synthetisized cDNAs were not fractionated through the gel filtration medium contained in the kit but instead through Size Sep 400 Spun Columns (Pharmacia).

Complete polynucleotide sequences

Apart from the first screening round when isolating the Hva7-1 clone, all the screens were carried out using the DIG system (all reagents and consumables from Boehringer Mannheim, in accordance with the instructions in "The DIG System User's Guide for Filter Hybridization", Boehringer Mannheim). The DNA probes employed were prepared by means of PCR using digoxigenin-labelled dUTP. The hybridizations were carried out at 42°C overnight in DIG Easy Hyb (Boehringer Mannheim). Labelled DNA was detected on nylon membranes by means of chemiluminescence (CDP-Star, Boehringer Mannheim) using X-ray films (Hyperfilm MP, Amersham). Initial partial sequencing of the isolated gene library plasmids was

carried out, for identification purposes, using T3 and T7 primers (ABI Prism Dye Terminator Cycle Sequencing Kit, ABI, using an ABI Prism 310 Genetic Analyzer). The complete polynucleotide sequences in Hva7-1, Hva7-2 and Da7 were determined, as a commissioned sequencing carried out by Qiagen, Hilden, by means of primer walking using cycle sequencing.

a. Isolating the Da7 clone

10^6 phages from a Drosophila melanogaster cDNA library in λ phages (Canton-S embryo, 2-14 hours, in Uni-ZAP XR vector, Stratagene) were screened using DIG-labelled 248-5 as the probe (in accordance with the manufacturer's (Stratagene) instructions). The maximum stringency when washing the filters was: 0.2 x SSC; 0.1% SDS; 42°C; 2 x 15 min. One clone (clone 432-1) was isolated whose insert had a size of 2940 bp (Da7, SEQ ID NO: 1). The largest open reading frame of this sequence begins at position 372 of the depicted sequence and ends at position 1822. The 770 amino acid polypeptide sequence which is deduced from this (SEQ ID NO: 2) encodes a protein having a calculated molecular weight of 87.01 kD.

b. Isolating the Hva7-1 clone

10^6 phages from the Heliothis virescens embryo cDNA library (library I) were included in the screening. The first of three screening rounds took place using α-^{32}P-labelled 228-11 DNA as the probe. The probe was hybridized to the filters in Quickhyb (Stratagene) at 68°C for one hour. The filters were then washed twice, for 15 min on each occasion, at room temperature in 2 x SSC; 0.1% SDS and twice, for 30 min on each occassion, at 42°C in 0.1xSSC; 0,1% SDS. Hybridized probe was detected by means of autoradiography, at -80°C overnight, using XR X-ray films (Kodak) and employing intensifying screens (Amersham). The two further screening rounds were carried out using the DIG System (Boehringer Mannheim).

The clone 241-5, which was isolated in this screen, contained an insert of 3630 bp. This insert (Hva7-1, SEQ ID NO: 3) possesses a longest open reading frame which begins at position 335 of the depicted nucleic acid sequence and ends at position 1821. The 496 amino acid polypeptide which is deduced from this (SEQ ID NO: 4) encodes a protein having a calculated molecular weight of 56.36 kD.

c. Isolating the Hva7-2 clone

10^6 phages from the Heliothis virescens ganglia cDNA library (library II) were included in the screening. Dig-labelled 228-8 DNA was used as the probe. The maximum stringency when washing the filters was: 0.1 x SSC; 0.1% SDS; 42°C; 2 x 15 min.

The clone 241-5, which was isolated in this screen, contained an insert of 3630 bp. This insert (Hva7-2, SEQ ID NO: 5) possesses a longest open reading frame which begins at position 95 of the depicted nucleic acid sequence and ends at position 1598. The 501 amino acids polypeptide which was deduced from this (SEQ ID NO: 6) encodes a protein having a calculated molecular weight of 56.71 kD.

Example 2

Generating the expression constructs

a. Da7

The sequence region from position 372 to position 2681 of SEQ ID NO: 1 was amplified by means of a polymerase chain reaction (PCR). Deoxyoligonucleotides having the sequences
GCGAATTCACCACCATGAAAAATGCACAACTG and
CGAGACAATAATATGTGGTGCCTCGAG were used for this. The Pfu polymerase from Stratagene was used as the DNA polymerase in accordance with the manufacturer's instructions. Following the amplification, the segment which had been generated was digested with the restriction endonucleases Eco RI and Xho I and cloned into a vector, i.e. pcDNA3.1/Zeo (Invitrogen), which had likewise been digested with Eco RI and Xho I.

b. Hva7-1

The sequence region from position 335 to position 1822 from SEQ ID NO: 3 was amplified by means of a polymerase chain reaction (PCR). Deoxyoligonucleotides having the sequences

GCAAGCTTACCACCATGGGAGGTAGAGCTAGACGCTCGCAC and GCCTCGAGCGACACCATGATGTGTGGCGC were used for this. The Pfu polymerase from Stratagene was used as the DNA polymerase in accordance with the manufacturer's instructions.

Following amplification, the generated segment was digested with the restriction endonucleases HindIII and Xho I and cloned into a vector, i.e. pcDNA3.1/Zeo (Invitrogen), which had likewise been digested with HindIII and Xho I.

c. Hva7-2

The sequence region from position 95 to position 1597 from SEQ ID NO: 5 was amplified by means of a polymerase chain reaction (PCR). Deoxyoligonucleotides having the sequences GCAAGCGCCGCTATGGCCCCTATGTTG and TTGCACGATGATATGCGGTGCCTCGAGCG were used for this. The Pfu polymerase from Stratagene was used as the DNA polymerase in accordance with the manufacturer's instructions. Following amplification, the generated segment was digested with the restriction endonucleases HindIII and Xho I and cloned into a vector, i.e. pcDNA3.1/Zeo (Invitrogen), which had likewise been digested with HindIII and Xho I.

d.Hva7-1 / $5HT_3$ and Hva7-2 / $5HT_3$ chimaeras

The region from position 335 to position 1036 from SEQ ID NO: 3 (Hva7-1/$5HT_3$ chimaera) and the region from position 95 to position 763 from SEQ ID NO: 5 (Hva7-2/$5HT_3$ chimaera) was in each case fused to the region from position 778 to position 1521 from the Mus musculus 5-HT_3 receptor cDNA (sequence in EMBL database: M774425) using the method of overlap extension (Jespersen et al. 1997). The two fragments were subsequently cloned into the pcDNA3.1/Zeo vector by means of TA cloning (Invitrogen, in accordance with the manufacturer's instructions). Constructs containing the correct orientation of the two fragments in the vector were identified by sequencing using the T7 primer (Invitrogen).

Cell culture and gene transfer

HEK293 cells, which express the α subunit of an L-type Ca channel (Zong et al. 1995, Stetzer et al. 1996), were cultured in Dulbecco's modified Eagle's medium and 10% foetal calf serum at 5% CO_2 and from 20°C to 37°C. FuGENE 6 (Boehringer Mannheim GmbH, Mannheim, Germany) was used for the gene transfer in accordance with the manufacturer's instructions. At from 24 h to 48 h after the gene transfer, the cells were sown at various densities in microtitre plates. Recombinantly altered cells were selected by growth in Dulbecco's modified Eagle's medium and 10% foetal calf serum and 150 - 500 µg/ml of Zeocin/ml over a period of from 3 to 4 weeks. Individual resistant clones were analyzed as described below.

Fura-2 measurements

The alterations in the intracellular calcium concentration were measured using Fura-2. A stock solution containing 2 mM Fura-2-acetoxy methyl ester (Sigma) in dimethyl sulphoxide (DMSO) was diluted to a final concentration of 5 - 10 µM in serum-free minimal essential medium (MEM, Gibco) containing 1% bovine serum albumin and 5 mM calcium chloride. The cells were incubated for from 45 to 60 min in this solution in a microtitre plate. The cells were then washed twice in Tyrode solution buffered with N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulphonic acid) (5 mM HEPES) (HEPES-buffered salt solution containing 130 mM NaCl, 5 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, 5 mM $NaHCO_3$, 10 mM glucose). 100 µl Tyrode buffer were added to the wells of the microtitre plate and the cells were illuminated alternately, under a fluorescence microscope (Nikon Diaphot), with light of 340 nm and 380 nm wavelength. A series of video images (exposure time per image 100 ms) were taken with pauses of 3 seconds and stored, as digitalized images, in an image analysis computer (Leica, Quantimet 570). After 8 images had been taken (measurement point 4.0 in Fig. 1), nicotine was added to a final concentration of 500 µM and the measurement series was continued. The fluorescence intensity of the cells when illuminating with light of 380 nm wavelength was divided by the corresponding intensity at 340 nm and in this way a ratio was formed which represents the relative increase in calcium concentration (Grynkiewicz et al. 1985).

References:

Bossy et al. (1988) Conservation of neural nicotinic acetylcholine receptors from Drosophila to vertebrate central nervous systems, EMBO J. 7, 611-618

Breer et al. (1987) Molecular properties and functions of insect acetylcholine receptors, J. Insect Physiol. 33, 771-790

Buckingham et al. (1997) Imidacloprid actions on insect neuronal acetylcholine receptors, J.Exp. Biol. 200, 2685-2692

Changeux et al. (1992) The functional architecture of the nicotinic acetylcholine receptor explored by affinity labelling and site-directed mutagenesis, Quarterly Review of Biophysics 25, 395-432

Claudio et al. (1983) Nucleotide and deduced amino acid sequences of Torpedo californica acetylcholine receptor g subunit, Proc. Natl. Acad. Sci. USA 80, 1111-1115

Devillers-Thiery et al. (1983) Complete mRNA coding sequence of the acetylcholine binding α-subunit of Torpedo marmorata acetylcholine receptor: a model for the transmembrane organization of the polypeptide chain, Proc. Natl. Acad. Sci. USA 80, 2067-2071

Elgoyhen et al. (1997) US Pat. No. 5,683,912

Eastham et al. (1998) Characterisation of a nicotinic acetylcholine receptor from the insect Manduca sexta, Eur. J. Neurosci 10, 879-889

Grynkiewicz et al. (1985) A new generation of Ca2+ indicators with greatly improved fluorescence properties, J. Biol. Chem. 260, 3440-3450

Hay et al. (1997), P element insertion-dependent gene activation in the Drosophila eye, Proceedings of The National Academy of Sciences of The United States of America 94 (10), 5195-5200

Hermans-Borgmeyer et al. (1986) Primary structure of a developmentally regulated nicotinic acetylcholine receptor protein from Drosophila EMBO J. 5, 1503-1508

Heinemann et al. (1997) US Pat. No 5,591,590

Jespersen et al. (1997) Efficient Non-PCR-Mediated Overlap Extension of PCR Fragments by Exonuclease "End Polishing", Biotechniques, 23, 48-52

Lindstrom et al. (1997) US Pat. No. 5,599,709

Marshall et al. (1990) Sequence and functional expression of a single a subunit of an insect nicotinic acetylcholine receptor, EMBO J. 9, 4391-4398

Noda et al. (1982), Primary structure of α-subunit precursor of Torpedo californica acetylcholine receptor deduced from cDNA sequence, Nature 299, 793-797

Noda et al. (1983a), Primary structures of β- and δ-subunit precursor of Torpedo californica acetylcholine receptor deduced from cDNA sequences, Nature 301, 251-255

Noda et al. (1983b), Structural homology of Torpedo californica acetylcholine receptor subunits, Nature 302, 528-532

Ortells et al. (1995), Evolutionary history of the ligand-gated ion-channel superfamily of receptors, Trends in Neurosience 18, 121-127

Plasterk (1996), The Tc1/mariner transposon family, Transposable Elements, Current Topics in Microbiology and Immunology 204, 125-143

Sambrook et al. (1989), Molecular Cloning, A Laboratory Manual, 2nd ed. Cold Spring Harbour Press

Sawruk et al. (1990a), EMBO J. 9, 2671-2677 Heterogeneity of Drosophila nicotinic acetylcholine receptors: SAD, a novel developmentally regulated α-subunit

Sawruk et al. (1990b), SBD, a novel structural subunit of the Drosophila nicotinic acetylcholine receptor, shares its genomic localization with two a-subunits, FEBS Lett. 273, 177-181

Schloß et al. (1988), Neuronal acetylcholine receptors of Drosophila: the ARD protein is a component of a high-affinity α-bungarotoxin binding complex, EMBO J. 7, 2889-2984

Stetzer et al. (1996) Stable expression in HEK-293 cells of the rat α3/β4 subtype of neuronal nicotinic acetylcholine receptor, FEBS Lett. 397, 39-44

Zong et al. (1995) On the regulation of the expressed L-type calcium channel by cAMP-dependent phosphorylation, Pflügers Arch. - Eur. J. Physiol. 430, 340-347.

SEQUENCE LISTING

(1) GENERAL INFORMATION:
(i) APPLICANT:
(A) NAME: Bayer Aktiengesellschaft
(B) STREET: Bayerwerk
(C) CITY: Leverkusen
(E) COUNTRY: Germany
(F) POSTAL CODE: D 51368

(ii) TITLE OF THE INVENTION: Nucleic acids which encode insect acetylcholine receptor subunits

(iii) NUMBER OF SEQUENCES: 6

(iv) COMPUTER-READABLE FORM:
(A) MEDIUM TYPE: floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2886 base pairs
(B) TYPE: nucleotide
(C) STRANDEDNESS: double strand
(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: cDNA of mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Drosophila melanogaster

(vii) IMMEDIATE SOURCE:
(B) CLONE(S): Da7

(ix) FEATURES:
(A) NAME/KEY: CDS
(B) POSITION: 372..2681

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GGCACGAGAA AAAGTTGTGG TATAAACTTT TATTGTAGGA AAACGCATAA AAATAATAGA 60

AAAACGCTCT TCGGGTTGTA AAGAAAATAA GAAGACAAAA GAAAGACATG AAAACGTTGC 120

AAACAATAAA GCATATACTT GCCATATTGA TATAAAGGGA AATCGTGAAA AGGCGGTGAA 180

AATTTCGTAA GATTAGTTGG TATTAAGGGC AGCCCATGCA CACAGCTAAA AAGGGAACTA 240

```
AAAAAACCCC GCACAGAACA ATGAAAGCTG CAGCAGCTGG ATAAGGCCGA CAAAACCGAA      300

AATTATATTA TTGTAATCTA GTAGAGAGCA GACAACATAT CCGCTGGCAA CAACCAACAC      360

CGAAAGAGAC T ATG AAA AAT GCA CAA CTG AAA CTG ACT GAA GTT GAC GAT        410
             Met Lys Asn Ala Gln Leu Lys Leu Thr Glu Val Asp Asp
               1               5                  10

GAT GAG CTG TGG CTG GCA GTA AGA TTA GCG CAC TGC AGC AGC AAC TTT        458
Asp Glu Leu Trp Leu Ala Val Arg Leu Ala His Cys Ser Ser Asn Phe
     15                  20                  25

AGC AGC AGT AGC AGC ACA AGA ACC ACC AGC AGC AAC CAG AGG CAC AAC        506
Ser Ser Ser Ser Ser Thr Arg Thr Thr Ser Ser Asn Gln Arg His Asn
 30                  35                  40                  45

CAG CAA CTC ACA ACA CTG CAA CCA AGG AGC TTA AGT ACA AAA CAC CAC        554
Gln Gln Leu Thr Thr Leu Gln Pro Arg Ser Leu Ser Thr Lys His His
                 50                  55                  60

AGC AAC ATT GCA AGC GAG CAG CAC AAT AGC CAG CAA CAG GAG CCA GCA        602
Ser Asn Ile Ala Ser Glu Gln His Asn Ser Gln Gln Gln Glu Pro Ala
           65                  70                  75

TCG AAG GAC GAG GAT GTA GCC AAC CAC GGT AGA AGC AAT GAC CAG CAG        650
Ser Lys Asp Glu Asp Val Ala Asn His Gly Arg Ser Asn Asp Gln Gln
      80                  85                  90

ACG CAT CTG CAA CAG CTA GAC AGC AGC AAC ATG TTG TCG CCA AAG ACA        698
Thr His Leu Gln Gln Leu Asp Ser Ser Asn Met Leu Ser Pro Lys Thr
 95                  100                 105

GCC GCA GCA GCA ACT GCT GCC GGC GAT GAA GCA ACA ACC CAA CAA CCA        746
Ala Ala Ala Ala Thr Ala Ala Gly Asp Glu Ala Thr Thr Gln Gln Pro
110                 115                 120                 125

ACA AAC ATA AGA CTG TGT GCA CGC AAG CGA CAA CGA TTG CGT CGC CGA        794
Thr Asn Ile Arg Leu Cys Ala Arg Lys Arg Gln Arg Leu Arg Arg Arg
                130                 135                 140

CGA AAA AGA AAA CCA GCA ACC CCA AAC GAA ACA GAT ATC AAG AAA CAA        842
Arg Lys Arg Lys Pro Ala Thr Pro Asn Glu Thr Asp Ile Lys Lys Gln
           145                 150                 155

CAG CAA CTT AGC ATG CCT CCC TTC AAA ACG CGC AAA TCC ACG GAC ACC        890
Gln Gln Leu Ser Met Pro Pro Phe Lys Thr Arg Lys Ser Thr Asp Thr
      160                 165                 170

TAC AGC ACA CCA GCA GCA ACA ACC AGC TGT CCG ACA GCC ACC TAC ATG        938
Tyr Ser Thr Pro Ala Ala Thr Thr Ser Cys Pro Thr Ala Thr Tyr Met
 175                 180                 185

CAA TGT CGA GCC AGC GAC AAT GAG TTC AGT ATT CCG ATA TCG AGA CAT        986
Gln Cys Arg Ala Ser Asp Asn Glu Phe Ser Ile Pro Ile Ser Arg His
190                 195                 200                 205

GAT AGA GTA TCC ACG GCC ACA TTC GCC TGG GTG TTG CAT GTG CTG CAG       1034
Asp Arg Val Ser Thr Ala Thr Phe Ala Trp Val Leu His Val Leu Gln
                210                 215                 220

GTG CTG CTC GTG TCG CTG CAA CAG TGG CAA CTT CAC GTG CAA CAG CGA       1082
Val Leu Leu Val Ser Leu Gln Gln Trp Gln Leu His Val Gln Gln Arg
           225                 230                 235

TCG GTG CTA CTG TTC AGA AGG ATC GCA GCG AGC ACC ATC GCC TTC ATT       1130
Ser Val Leu Leu Ph  Arg Arg Ile Ala Ala Ser Thr Ile Ala Phe Ile
      240                 245                 250
```

```
TCC TAT TTA GGC AGC TTT GCA GCG CAA CTG AAA AAT AGC AGC AGC AGC      1178
Ser Tyr Leu Gly Ser Phe Ala Ala Gln Leu Lys Asn Ser Ser Ser Ser
    255                 260                 265

AGT AGC AGC AGC AAC AGC AGC AAC AAC AGC AGC ACG CAA ATA TTA AAC      1226
Ser Ser Ser Ser Asn Ser Ser Asn Asn Ser Ser Thr Gln Ile Leu Asn
270                 275                 280                 285

GGA CTT AAT AAA CAC TCA TGG ATA TTT TTA TTG ATA TAT TTG AAT TTA      1274
Gly Leu Asn Lys His Ser Trp Ile Phe Leu Leu Ile Tyr Leu Asn Leu
                290                 295                 300

TCT GCT AAA GTT TGC CTA GCA GGA TAT CAT GAA AAG AGA CTG TTA CAC      1322
Ser Ala Lys Val Cys Leu Ala Gly Tyr His Glu Lys Arg Leu Leu His
            305                 310                 315

GAT CTT TTG GAT CCT TAT AAT ACA CTA GAA CGT CCC GTT CTC AAT GAA      1370
Asp Leu Leu Asp Pro Tyr Asn Thr Leu Glu Arg Pro Val Leu Asn Glu
        320                 325                 330

TCG GAC CCG TTA CAA TTA AGC TTT GGT TTA ACT TTA ATG CAA ATT ATC      1418
Ser Asp Pro Leu Gln Leu Ser Phe Gly Leu Thr Leu Met Gln Ile Ile
    335                 340                 345

GAT GTG GAC GAG AAA AAT CAA TTG CTA GTC ACT AAT GTG TGG TTA AAA      1466
Asp Val Asp Glu Lys Asn Gln Leu Leu Val Thr Asn Val Trp Leu Lys
350                 355                 360                 365

CTG GAG TGG AAC GAC ATG AAT CTC CGC TGG AAC ACC TCC GAC TAT GGC      1514
Leu Glu Trp Asn Asp Met Asn Leu Arg Trp Asn Thr Ser Asp Tyr Gly
                370                 375                 380

GGA GTT AAG GAT CTG CGA ATA CCG CCG CAT CGC ATC TGG AAG CCG GAC      1562
Gly Val Lys Asp Leu Arg Ile Pro Pro His Arg Ile Trp Lys Pro Asp
            385                 390                 395

GTG CTG ATG TAC AAC AGT GCG GAT GAG GGA TTT GAC GGC ACC TAC CAG      1610
Val Leu Met Tyr Asn Ser Ala Asp Glu Gly Phe Asp Gly Thr Tyr Gln
        400                 405                 410

ACG AAC GTG GTG GTG CGG AAC AAC GGC TCG TGT CTA TAC GTT CCG CCG      1658
Thr Asn Val Val Val Arg Asn Asn Gly Ser Cys Leu Tyr Val Pro Pro
    415                 420                 425

GGG ATC TTC AAG TCG ACG TGC AAG ATC GAC ATC ACG TGG TTC CCC TTC      1706
Gly Ile Phe Lys Ser Thr Cys Lys Ile Asp Ile Thr Trp Phe Pro Phe
430                 435                 440                 445

GAT GAC CAG CGG TGC GAG ATG AAG TTC GGC AGT TGG ACC TAC GAC GGA      1754
Asp Asp Gln Arg Cys Glu Met Lys Phe Gly Ser Trp Thr Tyr Asp Gly
                450                 455                 460

TTC CAG CTG GAT TTA CAA TTA CAA GAT GAA ACT GGC GGT GAT ATC AGC      1802
Phe Gln Leu Asp Leu Gln Leu Gln Asp Glu Thr Gly Gly Asp Ile Ser
            465                 470                 475

AGT TAC GTG CTC AAC GGC GAG TGG GAA CTA CTG GGT GTG CCC GGC AAA      1850
Ser Tyr Val Leu Asn Gly Glu Trp Glu Leu Leu Gly Val Pro Gly Lys
        480                 485                 490

CGT AAC GAG ATC TAT TAC AAC TGC TGC CCG GAA CCC TAT ATA GAC ATC      1898
Arg Asn Glu Ile Tyr Tyr Asn Cys Cys Pro Glu Pro Tyr Ile Asp Ile
    495                 500                 505

ACC TTC GCC ATC ATC ATC CGC CGA CGA ACA CTG TAC TAT TTC TTC AAC      1946
Thr Phe Ala Ile Ile Ile Arg Arg Arg Thr Leu Tyr Tyr Phe Ph  Asn
510                 515                 520                 525
```

```
CTG ATC ATA CCT TGT GTA CTG ATT GCC TCC ATG GCC TTG CTC GGA TTC      1994
Leu Ile Ile Pro Cys Val Leu Ile Ala Ser Met Ala Leu Leu Gly Phe
                530                 535                 540

ACC CTG CCG CCA GAT TCG GGT GAA AAA TTA TCG CTG GGT GTT ACC ATC      2042
Thr Leu Pro Pro Asp Ser Gly Glu Lys Leu Ser Leu Gly Val Thr Ile
            545                 550                 555

TTG CTC TCG CTG ACC GTG TTT CTG AAT ATG GTT GCC GAG ACA ATG CCG      2090
Leu Leu Ser Leu Thr Val Phe Leu Asn Met Val Ala Glu Thr Met Pro
        560                 565                 570

GCT ACT TCC GAT GCG GTG CCA TTG TGG ATA CGC ATC GTG TTT TTG TGC      2138
Ala Thr Ser Asp Ala Val Pro Leu Trp Ile Arg Ile Val Phe Leu Cys
    575                 580                 585

TGG CTG CCA TGG ATA TTG CGA ATG AGT CGC CCA GGA CGA CCG CTG ATC      2186
Trp Leu Pro Trp Ile Leu Arg Met Ser Arg Pro Gly Arg Pro Leu Ile
590                 595                 600                 605

CTA GAG TTC CCG ACC ACG CCC TGT TCG GAC ACA TCC TCC GAG CGG AAG      2234
Leu Glu Phe Pro Thr Thr Pro Cys Ser Asp Thr Ser Ser Glu Arg Lys
                610                 615                 620

CAC CAG ATA CTC TCC GAC GTT GAG CTG AAA GAG CGC TCG TCG AAA TCG      2282
His Gln Ile Leu Ser Asp Val Glu Leu Lys Glu Arg Ser Ser Lys Ser
            625                 630                 635

CTG CTG GCC AAC GTA CTA GAC ATC GAT GAT GAC TTC CGG CAC AAT TGT      2330
Leu Leu Ala Asn Val Leu Asp Ile Asp Asp Asp Phe Arg His Asn Cys
        640                 645                 650

CGC CCC ATG ACG CCC GGC GGA ACA CTG CCA CAC AAC CCG GCT TTC TAT      2378
Arg Pro Met Thr Pro Gly Gly Thr Leu Pro His Asn Pro Ala Phe Tyr
    655                 660                 665

CGC ACG GTT TAT GGA CAA GGC GAC GAT GGC AGC ATT GGG CCA ATT GGC      2426
Arg Thr Val Tyr Gly Gln Gly Asp Asp Gly Ser Ile Gly Pro Ile Gly
670                 675                 680                 685

AGC ACC CGA ATG CCG GAT GCG GTC ACC CAT CAT ACG TGC ATC AAA TCA      2474
Ser Thr Arg Met Pro Asp Ala Val Thr His His Thr Cys Ile Lys Ser
                690                 695                 700

TCA ACT GAA TAT GAA TTA GGT TTA ATC TTA AAG GAA ATT CGC TTT ATA      2522
Ser Thr Glu Tyr Glu Leu Gly Leu Ile Leu Lys Glu Ile Arg Phe Ile
            705                 710                 715

ACT GAT CAG CTA CGT AAA GAT GAC GAG TGC AAT GAC ATT GCC AAT GAT      2570
Thr Asp Gln Leu Arg Lys Asp Asp Glu Cys Asn Asp Ile Ala Asn Asp
        720                 725                 730

TGG AAA TTT GCA GCT ATG GTC GTT GAC AGA CTG TGC CTT ATC ATA TTC      2618
Trp Lys Phe Ala Ala Met Val Val Asp Arg Leu Cys Leu Ile Ile Phe
    735                 740                 745

ACA ATG TTC GCA ATA TTA GCC ACA ATA GCT GTA CTA CTA TCG GCA CCA      2666
Thr Met Phe Ala Ile Leu Ala Thr Ile Ala Val Leu Leu Ser Ala Pro
750                 755                 760                 765

CAT ATT ATT GTC TCG TAGCCATATG GGCGAGGTGG TTATTGTTAT TGGTTTTATT      2721
His Ile Ile Val Ser
            770

ATAAAATCAA TTTGTTAATT ATTAAATTAA TAACGAAACT CTTTAAGTAA ATTAAAACTA    2781

AAAAGACACT AAAAAAGCAC AAAAAAATAG GAAAATACAT GATAAAACCC ATGAACTAAA    2841
```

TAATACATCC AAGAAAAACC AAAACAAAAA AAAAAAAAAA AAAAA 2886

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 770 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Met Lys Asn Ala Gln Leu Lys Leu Thr Glu Val Asp Asp Asp Glu Leu
 1               5                  10                  15

Trp Leu Ala Val Arg Leu Ala His Cys Ser Ser Asn Phe Ser Ser Ser
            20                  25                  30

Ser Ser Thr Arg Thr Thr Ser Ser Asn Gln Arg His Asn Gln Gln Leu
        35                  40                  45

Thr Thr Leu Gln Pro Arg Ser Leu Ser Thr Lys His His Ser Asn Ile
    50                  55                  60

Ala Ser Glu Gln His Asn Ser Gln Gln Gln Glu Pro Ala Ser Lys Asp
65                  70                  75                  80

Glu Asp Val Ala Asn His Gly Arg Ser Asn Asp Gln Gln Thr His Leu
                85                  90                  95

Gln Gln Leu Asp Ser Ser Asn Met Leu Ser Pro Lys Thr Ala Ala Ala
            100                 105                 110

Ala Thr Ala Ala Gly Asp Glu Ala Thr Thr Gln Gln Pro Thr Asn Ile
        115                 120                 125

Arg Leu Cys Ala Arg Lys Arg Gln Arg Leu Arg Arg Arg Arg Lys Arg
    130                 135                 140

Lys Pro Ala Thr Pro Asn Glu Thr Asp Ile Lys Lys Gln Gln Gln Leu
145                 150                 155                 160

Ser Met Pro Pro Phe Lys Thr Arg Lys Ser Thr Asp Thr Tyr Ser Thr
                165                 170                 175

Pro Ala Ala Thr Thr Ser Cys Pro Thr Ala Thr Tyr Met Gln Cys Arg
            180                 185                 190

Ala Ser Asp Asn Glu Phe Ser Ile Pro Ile Ser Arg His Asp Arg Val
        195                 200                 205

Ser Thr Ala Thr Phe Ala Trp Val Leu His Val Leu Gln Val Leu Leu
    210                 215                 220

Val Ser Leu Gln Gln Trp Gln Leu His Val Gln Gln Arg Ser Val Leu
225                 230                 235                 240

Leu Phe Arg Arg Ile Ala Ala Ser Thr Ile Ala Phe Ile Ser Tyr Leu
                245                 250                 255

Gly Ser Phe Ala Ala Gln Leu Lys Asn Ser Ser Ser Ser Ser Ser Ser
            260                 265                 270

Ser Asn S r Ser Asn Asn Ser Ser Thr Gln Ile Leu Asn Gly Leu Asn
        275                 280                 285
```

```
Lys His Ser Trp Ile Phe Leu Leu Ile Tyr Leu Asn Leu Ser Ala Lys
    290                 295                 300

Val Cys Leu Ala Gly Tyr His Glu Lys Arg Leu Leu His Asp Leu Leu
305                 310                 315                 320

Asp Pro Tyr Asn Thr Leu Glu Arg Pro Val Leu Asn Glu Ser Asp Pro
                325                 330                 335

Leu Gln Leu Ser Phe Gly Leu Thr Leu Met Gln Ile Ile Asp Val Asp
            340                 345                 350

Glu Lys Asn Gln Leu Leu Val Thr Asn Val Trp Leu Lys Leu Glu Trp
        355                 360                 365

Asn Asp Met Asn Leu Arg Trp Asn Thr Ser Asp Tyr Gly Gly Val Lys
    370                 375                 380

Asp Leu Arg Ile Pro Pro His Arg Ile Trp Lys Pro Asp Val Leu Met
385                 390                 395                 400

Tyr Asn Ser Ala Asp Glu Gly Phe Asp Gly Thr Tyr Gln Thr Asn Val
                405                 410                 415

Val Val Arg Asn Asn Gly Ser Cys Leu Tyr Val Pro Pro Gly Ile Phe
            420                 425                 430

Lys Ser Thr Cys Lys Ile Asp Ile Thr Trp Phe Pro Phe Asp Asp Gln
        435                 440                 445

Arg Cys Glu Met Lys Phe Gly Ser Trp Thr Tyr Asp Gly Phe Gln Leu
    450                 455                 460

Asp Leu Gln Leu Gln Asp Glu Thr Gly Gly Asp Ile Ser Ser Tyr Val
465                 470                 475                 480

Leu Asn Gly Glu Trp Glu Leu Leu Gly Val Pro Gly Lys Arg Asn Glu
                485                 490                 495

Ile Tyr Tyr Asn Cys Cys Pro Glu Pro Tyr Ile Asp Ile Thr Phe Ala
            500                 505                 510

Ile Ile Ile Arg Arg Arg Thr Leu Tyr Tyr Phe Phe Asn Leu Ile Ile
        515                 520                 525

Pro Cys Val Leu Ile Ala Ser Met Ala Leu Leu Gly Phe Thr Leu Pro
    530                 535                 540

Pro Asp Ser Gly Glu Lys Leu Ser Leu Gly Val Thr Ile Leu Leu Ser
545                 550                 555                 560

Leu Thr Val Phe Leu Asn Met Val Ala Glu Thr Met Pro Ala Thr Ser
                565                 570                 575

Asp Ala Val Pro Leu Trp Ile Arg Ile Val Phe Leu Cys Trp Leu Pro
            580                 585                 590

Trp Ile Leu Arg Met Ser Arg Pro Gly Arg Pro Leu Ile Leu Glu Phe
        595                 600                 605

Pro Thr Thr Pro Cys Ser Asp Thr Ser Ser Glu Arg Lys His Gln Ile
    610                 615                 620

Leu Ser Asp Val Glu Leu Lys Glu Arg Ser Ser Lys Ser Leu Leu Ala
625                 630                 635                 640

Asn Val Leu Asp Ile Asp Asp Asp Phe Arg His Asn Cys Arg Pro Met
```

```
                645                 650                 655

Thr Pro Gly Gly Thr Leu Pro His Asn Pro Ala Phe Tyr Arg Thr Val
            660                 665                 670

Tyr Gly Gln Gly Asp Asp Gly Ser Ile Gly Pro Ile Gly Ser Thr Arg
        675                 680                 685

Met Pro Asp Ala Val Thr His His Thr Cys Ile Lys Ser Ser Thr Glu
    690                 695                 700

Tyr Glu Leu Gly Leu Ile Leu Lys Glu Ile Arg Phe Ile Thr Asp Gln
705                 710                 715                 720

Leu Arg Lys Asp Asp Glu Cys Asn Asp Ile Ala Asn Asp Trp Lys Phe
                725                 730                 735

Ala Ala Met Val Val Asp Arg Leu Cys Leu Ile Ile Phe Thr Met Phe
            740                 745                 750

Ala Ile Leu Ala Thr Ile Ala Val Leu Leu Ser Ala Pro His Ile Ile
        755                 760                 765

Val Ser
    770
```

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3700 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: double strand
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: cDNA of mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Heliothis virescens

(vii) IMMEDIATE SOURCE:
- (B) CLONE(S): Hva7-1

(ix) FEATURES:
- (A) NAME/KEY: CDS
- (B) POSITION: 335..1822

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
GGCACGAGCC GCTGCCCCAC GGTCGGCCGC ACTCCGCTGA ACAACAATGC TCAAAAACAC      60

GCCGTGACTC CACACACATC CCCTCGGCGC AGTAGGCGAT GTTTGAGGAT CGGACGGCAC     120

GCGTGGCCGT CGGCGAGCGG TCGTGAACAA GTTGCATACA TATGAAAACC GTAAAAAGAT     180

TGAATTTTAA GCCGATCGTG TTCGATAGAT CCTAATAGAG AAGCGGGAGT GCGGCGTTTG     240

GTAGGCGGGG GTCGAGTCGC GCGGTCGGGG GAAATGGCGC GGCGCGGGGC GGCGGCGGCG     300

GCGGCGCGCG GCGCGGCGGC GTCGCGGCGC TGAC ATG GGC GGG CGG GCG CGC          352
                                      Met Gly Gly Arg Ala Arg
                                                  775
```

```
CGC TCG CAC TTG GCG GCG CCC GCG GGC CTG CTG CTG CTG CTG TGC CTG        400
Arg Ser His Leu Ala Ala Pro Ala Gly Leu Leu Leu Leu Leu Cys Leu
            780                 785                 790

CTC TGG CCG AGG GGG GCA CGC TGC GGG TAC CAC GAG AAG CGG CTA CTG        448
Leu Trp Pro Arg Gly Ala Arg Cys Gly Tyr His Glu Lys Arg Leu Leu
        795                 800                 805

CAC CAC CTA TTG GAC CAC TAC AAC GTA CTG GAG AGG CCC GTC GTC AAC        496
His His Leu Leu Asp His Tyr Asn Val Leu Glu Arg Pro Val Val Asn
    810                 815                 820

GAG AGC GAC CCG CTG CAG CTC TCC TTC GGC CTC ACG CTC ATG CAG ATC        544
Glu Ser Asp Pro Leu Gln Leu Ser Phe Gly Leu Thr Leu Met Gln Ile
825                 830                 835                 840

ATC GAC GTG GAC GAG AAG AAC CAG CTT TTA ATA ACA AAC ATC TGG CTA        592
Ile Asp Val Asp Glu Lys Asn Gln Leu Leu Ile Thr Asn Ile Trp Leu
                845                 850                 855

AAA CTA GAG TGG AAT GAT ATG AAC TTG AGG TGG AAC ACT TCA GAT TTC        640
Lys Leu Glu Trp Asn Asp Met Asn Leu Arg Trp Asn Thr Ser Asp Phe
            860                 865                 870

GGC GGG GTC AAA GAT TTA AGA GTG CCA CCC CAC AGA CTA TGG AAA CCA        688
Gly Gly Val Lys Asp Leu Arg Val Pro Pro His Arg Leu Trp Lys Pro
        875                 880                 885

GAC GTC CTT ATG TAC AAC AGC GCG GAC GAA GGG TTC GAC AGC ACG TAT        736
Asp Val Leu Met Tyr Asn Ser Ala Asp Glu Gly Phe Asp Ser Thr Tyr
    890                 895                 900

CCA ACG AAC GTG GTG GTG CGG AAC AAC GGC TCG TGT CTG TAC GTG CCG        784
Pro Thr Asn Val Val Val Arg Asn Asn Gly Ser Cys Leu Tyr Val Pro
905                 910                 915                 920

CCC GGC ATC TTC AAG AGC ACC TGC AAG ATC GAC ATC ACC TGG TTC CCC        832
Pro Gly Ile Phe Lys Ser Thr Cys Lys Ile Asp Ile Thr Trp Phe Pro
                925                 930                 935

TTC GAC GAC CAA CGA TGC GAG ATG AAG TTT GGC AGC TGG ACT TAT GAT        880
Phe Asp Asp Gln Arg Cys Glu Met Lys Phe Gly Ser Trp Thr Tyr Asp
            940                 945                 950

GGT TAT CAG TTG GAT CTA CAA CTA CAG GAT GAA GGG GGC GGA GAT ATA        928
Gly Tyr Gln Leu Asp Leu Gln Leu Gln Asp Glu Gly Gly Gly Asp Ile
        955                 960                 965

AGC AGT TTT GTC ACG AAT GGC GAA TGG GAG TTA ATA GGA GTC CCC GGC        976
Ser Ser Phe Val Thr Asn Gly Glu Trp Glu Leu Ile Gly Val Pro Gly
    970                 975                 980

AAG CGC AAC GAG ATC TAC TAC AAC TGT TGT CCG GAG CCA TAC ATC GAC       1024
Lys Arg Asn Glu Ile Tyr Tyr Asn Cys Cys Pro Glu Pro Tyr Ile Asp
985                 990                 995                1000

ATC ACG TTT GCG GTG GTG ATC CGG AGG AAA ACG CTC TAC TAC TTC TTC       1072
Ile Thr Phe Ala Val Val Ile Arg Arg Lys Thr Leu Tyr Tyr Phe Phe
                1005                1010                1015

AAT CTG ATC GTG CCC TGC GTG CTC ATC GCC TCC ATG GCT CTA TTG GGG       1120
Asn Leu Ile Val Pro Cys Val Leu Ile Ala Ser Met Ala Leu Leu Gly
            1020                1025                1030

TTC ACC TTG CCT CCA GAC TCC GGA GAA AAG TTG TCT TTA GGT GTG ACG       1168
Phe Thr Leu Pro Pro Asp Ser Gly Glu Lys Leu Ser Leu Gly Val Thr
        1035                1040                1045
```

```
ATA TTA CTG TCG TTG ACG GTG TTC CTC AAC ATG GTG GCG GAG ACG ATG      1216
Ile Leu Leu Ser Leu Thr Val Phe Leu Asn Met Val Ala Glu Thr Met
    1050                1055                1060

CCA GCG ACG TCG GAC GCC GTG CCC TTG CTC GGC ACC TAC TTC AAC TGC      1264
Pro Ala Thr Ser Asp Ala Val Pro Leu Leu Gly Thr Tyr Phe Asn Cys
1065                1070                1075                1080

ATC ATG TTC ATG GTG GCT TCC TCC GTC GTC TCC ACC ATA CTG ATC CTC      1312
Ile Met Phe Met Val Ala Ser Ser Val Val Ser Thr Ile Leu Ile Leu
                1085                1090                1095

AAC TAC CAC CAC CGG CAC GCA GAC ACT CAC GAA ATG AGT GAT TGG ATT      1360
Asn Tyr His His Arg His Ala Asp Thr His Glu Met Ser Asp Trp Ile
            1100                1105                1110

CGT TGC GTG TTC CTT TAT TGG CTG CCG TGG GTG CTG CGC ATG TCA CGG      1408
Arg Cys Val Phe Leu Tyr Trp Leu Pro Trp Val Leu Arg Met Ser Arg
        1115                1120                1125

CCC GGC TCG GCG ACG ACG CCG CCG CCG GCG CGC GTA CCT CCG CCG CCG      1456
Pro Gly Ser Ala Thr Thr Pro Pro Pro Ala Arg Val Pro Pro Pro Pro
    1130                1135                1140

GAC CTG GAG CTG CGC GAG CGC TCC TCC AAG TCG CTC CTA GCG AAC GTG      1504
Asp Leu Glu Leu Arg Glu Arg Ser Ser Lys Ser Leu Leu Ala Asn Val
1145                1150                1155                1160

CTC GAC ATC GAT GAC GAC TTC CGC CAC CCG CAA GCG CAG CAG CCG CAA      1552
Leu Asp Ile Asp Asp Asp Phe Arg His Pro Gln Ala Gln Gln Pro Gln
                1165                1170                1175

TGC TGC CGA TAC TAC AGG GGG GGT GAG GAG AAT GGC GCG GGG TTG GCG      1600
Cys Cys Arg Tyr Tyr Arg Gly Gly Glu Glu Asn Gly Ala Gly Leu Ala
            1180                1185                1190

GCG CAC AGT TGC TTC GGT GTC GAC TAC GAG CTC TCC CTC ATT CTG AAG      1648
Ala His Ser Cys Phe Gly Val Asp Tyr Glu Leu Ser Leu Ile Leu Lys
        1195                1200                1205

GAG ATT AGA GTC ATC ACA GAT CAG ATG CGC AAG GAC GAC GAA GAT GCG      1696
Glu Ile Arg Val Ile Thr Asp Gln Met Arg Lys Asp Asp Glu Asp Ala
    1210                1215                1220

GAC ATT TCG CGC GAC TGG AAG TTC GCC GCC ATG GTC GTG GAC AGA CTG      1744
Asp Ile Ser Arg Asp Trp Lys Phe Ala Ala Met Val Val Asp Arg Leu
1225                1230                1235                1240

TGC CTT ATT ATC TTT ACC CTG TTC ACA ATC ATC GCC ACG CTA GCC GTG      1792
Cys Leu Ile Ile Phe Thr Leu Phe Thr Ile Ile Ala Thr Leu Ala Val
                1245                1250                1255

CTG CTG TCC GCG CCA CAC ATC ATG GTG TCG TAGCGACCCG CCCGCTTGCG        1842
Leu Leu Ser Ala Pro His Ile Met Val Ser
            1260                1265

GATACGCATG CGAAAAGTTC TGTGATACCG CGAATATTTG TTAAGTTGTG ATGAGCGAAG    1902

TGGCGCGGAC GGTGACGCCG CGGCGTCGGA GTTGCCGCCG CCTGCCTCGC CGCCCGCGCC    1962

CCCCTGTAGA CATAAGTTAC CGCTGACTGC CAACCCTGTA CGTTCAACAA ATAACTGCCC    2022

ATCCGACTAA CGTCTTTTAT CCCCTTGAAA AATTCAGCGA TTGTGTACCC CTTTCTTCCA    2082

AGAATACAAT GACAAATGGT CGTCACGCTC AGTGGAATCA ATCCCGTACT CTTCGCCCGA    2142

TATTTCCCTT AGGGTATGTC ACGAGTTTGA ATGAGCGGTT CCGTATCAGA CGTTCCGTCC    2202
```

CCGGAACGGT CGTCCCCTGC GATAAAGTGG CAGTACGTGC TATACAGGCA CTTAAGGCCG 2262

CCACGCCACG GCGCCGCGGT GCGCTCGGGC CGCGAACCCG CGACCCTCAC CGCTGCAAGT 2322

GGCCACCCAC TAGACAAGAC TGCGGCAGAA AATATTTGCA CAAAAACGTC TTCCTTCTTA 2382

CCGATGAACG ACCTGATTCG CATTTAAAAT TAAACTTTGT TAGAACTTCT TCGATTCTTG 2442

AAATCTATTG TACAGTTTAG AGTTTGGGCG GTGAAACAAT GGCCCTTTGT TTCCTTCTTG 2502

TTCGATTCCA TGAATCGTGG TTATAATCCC TAGTTTTATT TTCGGATATA TTTGTGTCAG 2562

TAGCTAGTAT AGAACTTTAC AAACAATGTT GATTCAATTG GTACAGGTTG TGATATGCCT 2622

CGTTGTGAAC GGGTCCGATA TTGTTATAAA TGGTAAAATA CCCATGGCTA TAGCTTAATA 2682

AATCGTTCGT TAAAAGTTGT AGTTAAACAA ATATTATTTT AATAAAGTCA TATCTGGGTC 2742

TTCCGGAACG ACTTTTACAA ATAATTAAAT TACATATTAA TATCACGTTT GTACTTCTTT 2802

CCATACAGTT ACAGTAATTC GTATGCTGAA AATAATATTA GCTTGTAAAA TTTTCTTCTT 2862

CGAAAATTTA TTCAAACAGA TGCGACCATC GTTTCAAACA TTTACATGTA ATATAGAACT 2922

CATTTTATAA GATATACAAC ATTTTATAAG TACAAGAAGT TGTAACATGA ACCGGTTTTT 2982

CGTTACATAG AGGGTATAAC ACAAAGGTGC CTACATATTG ACAGATGCGA AGCACGATCA 3042

GTTGATAAGC ACAGGTACAC TATATCCTGA CATCCGACAG TCCTGCCGCT CGTCTGCCAC 3102

ACTCGGAAAC ATTCGACAGT TCAGTTTACT GCTCCGCCAT CATCGATTGT TAAGTTTGTT 3162

GTTCTAACTC ATCGCATTCA TTTCATTCAA AAACATTGTA AACCTCTCAA GGGGAAAACG 3222

TGTTGTAAAC AGTGAGAGTG CGCGGGTACA ACCGACACGC GAATGTACCC TCGCAAGGCT 3282

CCTGTAATGT TTTCCTCTTC CGAGGTGTTG CTGAGAGTAA TCTTAGACGG TCCGATGGAA 3342

GTTGCGGACC GGATATGATT ACAAGTCAAT GTTTTTAAGT CATCCGTTTA TTTATTGTTA 3402

TATCTTCTTA CCATTCGCTA GAGGTTGTGT GACGACCCGG ACGGTGGGCG CCGCAACCCG 3462

CACACGCGGG GTTCCATCTT TGTATTAGAT GGAAGTTGTG CGGCATCTCT CCGTCGGCAA 3522

TGGGACAACC CGTTGTCCCC AACATTTGTT CAATTGTTAG GGTTAACTCT GAATTGCACT 3582

TTGTTTATTA AATATAAACG AATGAAACAA AAAAAAAAAA AAAAAACTCG AGAGTACTTC 3642

TAGAGCGGCC GCGGGCCCAT CGATTTTCCA CCCGGGTGGG GTACCAGTAA GTGTACCC 3700

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 496 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4

Met Gly Gly Arg Ala Arg Arg Ser His Leu Ala Ala Pro Ala Gly Leu
1 5 10 15

Leu Leu Leu Leu Cys Leu Leu Trp Pro Arg Gly Ala Arg Cys Gly Tyr
20 25 30

```
His Glu Lys Arg Leu Leu His His Leu Leu Asp His Tyr Asn Val Leu
          35                  40                  45

Glu Arg Pro Val Val Asn Glu Ser Asp Pro Leu Gln Leu Ser Phe Gly
     50                  55                  60

Leu Thr Leu Met Gln Ile Ile Asp Val Asp Glu Lys Asn Gln Leu Leu
 65                  70                  75                  80

Ile Thr Asn Ile Trp Leu Lys Leu Glu Trp Asn Asp Met Asn Leu Arg
                85                  90                  95

Trp Asn Thr Ser Asp Phe Gly Gly Val Lys Asp Leu Arg Val Pro Pro
            100                 105                 110

His Arg Leu Trp Lys Pro Asp Val Leu Met Tyr Asn Ser Ala Asp Glu
        115                 120                 125

Gly Phe Asp Ser Thr Tyr Pro Thr Asn Val Val Val Arg Asn Asn Gly
    130                 135                 140

Ser Cys Leu Tyr Val Pro Pro Gly Ile Phe Lys Ser Thr Cys Lys Ile
145                 150                 155                 160

Asp Ile Thr Trp Phe Pro Phe Asp Asp Gln Arg Cys Glu Met Lys Phe
                165                 170                 175

Gly Ser Trp Thr Tyr Asp Gly Tyr Gln Leu Asp Leu Gln Leu Gln Asp
            180                 185                 190

Glu Gly Gly Gly Asp Ile Ser Ser Phe Val Thr Asn Gly Glu Trp Glu
        195                 200                 205

Leu Ile Gly Val Pro Gly Lys Arg Asn Glu Ile Tyr Tyr Asn Cys Cys
    210                 215                 220

Pro Glu Pro Tyr Ile Asp Ile Thr Phe Ala Val Val Ile Arg Arg Lys
225                 230                 235                 240

Thr Leu Tyr Tyr Phe Phe Asn Leu Ile Val Pro Cys Val Leu Ile Ala
                245                 250                 255

Ser Met Ala Leu Leu Gly Phe Thr Leu Pro Pro Asp Ser Gly Glu Lys
            260                 265                 270

Leu Ser Leu Gly Val Thr Ile Leu Leu Ser Leu Thr Val Phe Leu Asn
        275                 280                 285

Met Val Ala Glu Thr Met Pro Ala Thr Ser Asp Ala Val Pro Leu Leu
    290                 295                 300

Gly Thr Tyr Phe Asn Cys Ile Met Phe Met Val Ala Ser Ser Val Val
305                 310                 315                 320

Ser Thr Ile Leu Ile Leu Asn Tyr His His Arg His Ala Asp Thr His
                325                 330                 335

Glu Met Ser Asp Trp Ile Arg Cys Val Phe Leu Tyr Trp Leu Pro Trp
            340                 345                 350

Val Leu Arg Met Ser Arg Pro Gly Ser Ala Thr Thr Pro Pro Pro Ala
        355                 360                 365

Arg Val Pro Pro Pro Pro Asp Leu Glu Leu Arg Glu Arg Ser Ser Lys
    370                 375                 380

Ser Leu Leu Ala Asn Val Leu Asp Ile Asp Asp Asp Phe Arg His Pro
```

```
385                 390                 395                 400

Gln Ala Gln Gln Pro Gln Cys Cys Arg Tyr Tyr Arg Gly Gly Glu Glu
                405                 410                 415

Asn Gly Ala Gly Leu Ala Ala His Ser Cys Phe Gly Val Asp Tyr Glu
            420                 425                 430

Leu Ser Leu Ile Leu Lys Glu Ile Arg Val Ile Thr Asp Gln Met Arg
        435                 440                 445

Lys Asp Asp Glu Asp Ala Asp Ile Ser Arg Asp Trp Lys Phe Ala Ala
    450                 455                 460

Met Val Val Asp Arg Leu Cys Leu Ile Ile Phe Thr Leu Phe Thr Ile
465                 470                 475                 480

Ile Ala Thr Leu Ala Val Leu Leu Ser Ala Pro His Ile Met Val Ser
                485                 490                 495
```

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3109 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: double strand
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: cDNA of mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Heliothis virescens

(vii) IMMEDIATE SOURCE:
- (B) CLONE(S): Hva7-2

(ix) FEATURES:
- (A) NAME/KEY: CDS
- (B) POSITION: 95..1597

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```
GGCACGAGCC GGCCGCACGT TGTCCCAGGC CGCATGAGCG CGCCGGCGTG CTAGCGCAGC      60

GTGCGCGGGT GTGGTATGCC CGCGCGTCGC CGCT ATG GCC CCT ATG TTG GCG          112
                                        Met Ala Pro Met Leu Ala
                                                    500

GCC TTG GCG CTG CTG GCT TTG CTG CCC GTA TCG GAG CAA GGT CCT CAC        160
Ala Leu Ala Leu Leu Ala Leu Leu Pro Val Ser Glu Gln Gly Pro His
        505                 510                 515

GAG AAG AGA CTC CTG AAC GCG TTG CTG GCG AAC TAC AAC ACC CTG GAG        208
Glu Lys Arg Leu Leu Asn Ala Leu Leu Ala Asn Tyr Asn Thr Leu Glu
    520                 525                 530

CGA CCG GTG GCC AAC GAG AGC GAA CCG CTA GAG GTC AGG TTC GGC TTG        256
Arg Pro Val Ala Asn Glu Ser Glu Pro Leu Glu Val Arg Phe Gly Leu
535                 540                 545                 550

ACC TTG CAG CAA ATC ATT GAC GTG GAC GAG AAG AAT CAA CTA CTT ATA        304
```

```
Thr Leu Gln Gln Ile Ile Asp Val Asp Glu Lys Asn Gln Leu Leu Ile
                555                 560                 565

ACC AAT ATA TGG CTG TCG TTG GAG TGG AAT GAC TAC AAC CTG AGG TGG      352
Thr Asn Ile Trp Leu Ser Leu Glu Trp Asn Asp Tyr Asn Leu Arg Trp
            570                 575                 580

AAC GAC AGC GAG TAT GGC GGG GTC AAG GAC CTC AGG ATC ACG CCC AAC      400
Asn Asp Ser Glu Tyr Gly Gly Val Lys Asp Leu Arg Ile Thr Pro Asn
        585                 590                 595

AAG TTG TGG AAG CCG GAC GTC CTT ATG TAT AAT AGT GCT GAC GAG GGT      448
Lys Leu Trp Lys Pro Asp Val Leu Met Tyr Asn Ser Ala Asp Glu Gly
    600                 605                 610

TTT GAC GGG ACC TAC CAG ACC AAC GTG GTG GTC AGA AGC GGC GGC AGT      496
Phe Asp Gly Thr Tyr Gln Thr Asn Val Val Val Arg Ser Gly Gly Ser
615                 620                 625                 630

TGC CTG TAC GTG CCA CCT GGC ATA TTC AAG AGC ACA TGC AAG ATG GAC      544
Cys Leu Tyr Val Pro Pro Gly Ile Phe Lys Ser Thr Cys Lys Met Asp
                635                 640                 645

ATC GCG TGG TTT CCC TTC GAC GAC CAA CAC TGT GAT ATG AAG TTC GGT      592
Ile Ala Trp Phe Pro Phe Asp Asp Gln His Cys Asp Met Lys Phe Gly
            650                 655                 660

AGC TGG ACA TAT GAC GGC AAT CAG TTG GAT CTG GTG CTA AAA GAT GAG      640
Ser Trp Thr Tyr Asp Gly Asn Gln Leu Asp Leu Val Leu Lys Asp Glu
        665                 670                 675

GCA GGC GGC GAT CTA TCG GAC TTC ATA ACA AAT GGG GAG TGG TAT CTA      688
Ala Gly Gly Asp Leu Ser Asp Phe Ile Thr Asn Gly Glu Trp Tyr Leu
    680                 685                 690

ATA GGA ATG CCA GGC AAA AAG AAC ACA ATA ACA TAC GCG TGC TGC CCC      736
Ile Gly Met Pro Gly Lys Lys Asn Thr Ile Thr Tyr Ala Cys Cys Pro
695                 700                 705                 710

GAG CCC TAC GTG GAC GTC ACC TTC ACC ATC ATG ATA AGA AGA CGA ACC      784
Glu Pro Tyr Val Asp Val Thr Phe Thr Ile Met Ile Arg Arg Arg Thr
                715                 720                 725

TTG TAC TAC TTC TTC AAC CTG ATC GTC CCG TGC GTG CTG ATC TCA TCG      832
Leu Tyr Tyr Phe Phe Asn Leu Ile Val Pro Cys Val Leu Ile Ser Ser
            730                 735                 740

ATG GCA CTC CTC GGC TTC ACA CTG CCA CCA GAC TCC GGA GAG AAA CTC      880
Met Ala Leu Leu Gly Phe Thr Leu Pro Pro Asp Ser Gly Glu Lys Leu
        745                 750                 755

ACA CTT GGA GTC ACT ATT CTT CTA TCG CTG ACG GTG TTC CTC AAC CTG      928
Thr Leu Gly Val Thr Ile Leu Leu Ser Leu Thr Val Phe Leu Asn Leu
    760                 765                 770

GTA GCC GAG ACC CTG CCA CAG GTC TCC GAC GCT ATC CCC CTG TTA GGG      976
Val Ala Glu Thr Leu Pro Gln Val Ser Asp Ala Ile Pro Leu Leu Gly
775                 780                 785                 790

ACG TAC TTC AAT TGC ATC ATG TTC ATG GTA GCG TCG TCT GTG GTA CTG     1024
Thr Tyr Phe Asn Cys Ile Met Phe Met Val Ala Ser Ser Val Val Leu
                795                 800                 805

ACT GTG GTG GTA CTC AAT TAC CAC CAT CGA ACA GCT GAT ATA CAT GAA     1072
Thr Val Val Val Leu Asn Tyr His His Arg Thr Ala Asp Ile His Glu
            810                 815                 820
```

```
ATG CCA CAG TGG ATA AAA TCA GTA TTC CTA CAA TGG TTG CCA TGG ATA        1120
Met Pro Gln Trp Ile Lys Ser Val Phe Leu Gln Trp Leu Pro Trp Il
        825                 830                 835

CTG CGA ATG TCG AGG CCA GGG AAG AAG ATC ACC AGG AAG ACT ATA ATG        1168
Leu Arg Met Ser Arg Pro Gly Lys Lys Ile Thr Arg Lys Thr Ile Met
    840                 845                 850

ATG AAC ACG AGG ATG AGG GAG CTG GAA CTG AAG GAG AGG TCG TCG AAG        1216
Met Asn Thr Arg Met Arg Glu Leu Glu Leu Lys Glu Arg Ser Ser Lys
855                 860                 865                 870

TCC TTG CTG GCG AAT GTT CTA GAT ATT GAT GAT GAC TTC AGA CAC GGC        1264
Ser Leu Leu Ala Asn Val Leu Asp Ile Asp Asp Asp Phe Arg His Gly
                875                 880                 885

CCT CCG CCT CCT AAC AGT ACT GCC TCG ACC GGG AAT TTG GGA CCT GGG        1312
Pro Pro Pro Pro Asn Ser Thr Ala Ser Thr Gly Asn Leu Gly Pro Gly
            890                 895                 900

TGC TCA ATA TTC CGC ACG GAT TTC CGT CGG TCG TTC GTC CGT CCG TCC        1360
Cys Ser Ile Phe Arg Thr Asp Phe Arg Arg Ser Phe Val Arg Pro Ser
        905                 910                 915

ACG ATG GAA GAC GTG GGC GGC GGG CTG GGT AGC CAC CAT CGC GAG CTG        1408
Thr Met Glu Asp Val Gly Gly Gly Leu Gly Ser His His Arg Glu Leu
    920                 925                 930

CAC CTC ATA CTG AGA GAG CTG CAG TTC ATC ACG GCC AGG ATG AAG AAG        1456
His Leu Ile Leu Arg Glu Leu Gln Phe Ile Thr Ala Arg Met Lys Lys
935                 940                 945                 950

GCT GAT GAG GAA GCC GAG CTG ATC AGC GAC TGG AAG TTT GCT GCG ATG        1504
Ala Asp Glu Glu Ala Glu Leu Ile Ser Asp Trp Lys Phe Ala Ala Met
                955                 960                 965

GTT GTT GAT AGG TTT TGC CTG TTC GTG TTC ACA CTT TTC ACA ATC ATC        1552
Val Val Asp Arg Phe Cys Leu Phe Val Phe Thr Leu Phe Thr Ile Ile
            970                 975                 980

GCG ACA GTA GCT GTC CTG TTA TCG GCA CCG CAT ATC ATC GTG CAA            1597
Ala Thr Val Ala Val Leu Leu Ser Ala Pro His Ile Ile Val Gln
        985                 990                 995

TGAACCAACC ACTGAGCCGG CAACTCCGGC GCATGAATGA GAGAAATAAT TATTAGATCG    1657

CCGATTTGTA ATTATAATTG ATAATGTAAT TAAATTAAAT ACGTGGTTGA AACGCACACG    1717

TCTCCATAAC AAAGTCTTAA GACATTAAAT TATGATAAAT TTACATATTG TAGTTAAGTC    1777

GAGTGTTGAT GGAAATTTTA GCCGGCGCAA GGAGTTTCGT GAAGGTCTGT ATATATTTTT    1837

TCTTATTGTT GTATATTGTA TCGTTGTTCA TGTTTTCTTT CAGGAAGTGA GCTTTGTACT    1897

GTTTGTTTCT TCGATGGCAG GTGCACTTCA GTTCAGGCTG AAATTTCCAT TAACATTTAT    1957

TTAAACAAAT GTGATGTTGA CTAGGATGTT ATACAGATAA ATGTTGACGT GTATAATTTG    2017

TTAAAATAAA CAATATTAAT TACTATTACT AAACGATATT ATAAACGAAG TACTAACGAG    2077

GGTTACTTTA ATGGGAAGAA CGCTAAGCTG GCACAGAGTT GCATTAATTT GAAAAAAGAA    2137

ATTACGGAAA AAAGTTTATT GAAAATTGAA CTTTTTGGAA GGAAAGTAAC GTTTGATCAA    2197

AAAAGTTTGT AAAACGAAAG TTCGGTTCTG CGCCAATACT GGAATTAAAA TTCTCGTAAA    2257

TATTAGGGAA AAGAAGGTCC TTTAAAACAA AAGATTTGAA CCGGCATCCT TTTTACAAGT    2317
```

AATGAGGGAT CACAGATGAT GACAAAAAAC CTTAGGGTAT ATAAGTAATG TACATAATGG 2377

ATCAAATATC GGTAGAGTCA AGAATAGTTA ACGATTTAAG ATTATTCCAT TCGATATTAA 2437

AATTCGATTA GCGATTGTCG CTGCGTCTAC TTTGATACAT ATCGATTTGA ATCGATATTG 2497

TATAAATTTA GATAGATCGG ACATTAGTAA TGAGTATGGA CGTTTTAATT TTTAAAAAAG 2557

AATGTACTAC GAAGATTAAA TCCAGGAATT GTTAAACAGT TATGGAATTG ATAAGAAATC 2617

AACAATTAAT ACGGAACCAA AGGTAGACTA GGTGTAGCAT CAGGAGATTG AATTAAAACA 2677

TAAATTAGGA CCGACTTAAA TGGAACTTGC GAGTGTATTG ATAACTTTTT AATTTAAAAA 2737

CTCATTGTCG ATTAAATGGA GAATAACTTT TGATCTCTCG TATCGATAAA TGCTCACTTA 2797

ACTATCGATA GCGTAATATT ATAACTGTTA GTATATCGAT ATGGGAGTAA GTCACTAGCA 2857

TCAGAAATAG TCATTAATTA GGAATCGGTT TGTGTTAATG TTATGCTTAG CGAAAATATT 2917

ACAATGCTGT TGATATCACT AACCATCACG TAACCATATT GATAAAATGT AAATACAGAA 2977

TATTGCGGTG TGTATTTGTA TATAAATTTT AGAAAAAAAA AAAAAAAAAA AACTCGAGAG 3037

TACTTCTAGA GCGGCCGCGG GCCCATCGAT TTTCCACCCG GGTGGGGTAC CAGGTAAGTG 3097

TACCCAATTC GC 3109

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 501 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```
Met Ala Pro Met Leu Ala Ala Leu Ala Leu Leu Ala Leu Leu Pro Val
 1               5                  10                  15

Ser Glu Gln Gly Pro His Glu Lys Arg Leu Leu Asn Ala Leu Leu Ala
            20                  25                  30

Asn Tyr Asn Thr Leu Glu Arg Pro Val Ala Asn Glu Ser Glu Pro Leu
        35                  40                  45

Glu Val Arg Phe Gly Leu Thr Leu Gln Gln Ile Ile Asp Val Asp Glu
    50                  55                  60

Lys Asn Gln Leu Leu Ile Thr Asn Ile Trp Leu Ser Leu Glu Trp Asn
65                  70                  75                  80

Asp Tyr Asn Leu Arg Trp Asn Asp Ser Glu Tyr Gly Gly Val Lys Asp
                85                  90                  95

Leu Arg Ile Thr Pro Asn Lys Leu Trp Lys Pro Asp Val Leu Met Tyr
            100                 105                 110

Asn Ser Ala Asp Glu Gly Phe Asp Gly Thr Tyr Gln Thr Asn Val Val
        115                 120                 125

Val Arg Ser Gly Gly Ser Cys Leu Tyr Val Pro Pro Gly Ile Phe Lys
    130                 135                 140

Ser Thr Cys Lys Met Asp Ile Ala Trp Phe Pro Phe Asp Asp Gln His
```

145 150 155 160

Cys Asp Met Lys Phe Gly Ser Trp Thr Tyr Asp Gly Asn Gln Leu Asp
165 170 175

Leu Val Leu Lys Asp Glu Ala Gly Gly Asp Leu Ser Asp Phe Ile Thr
180 185 190

Asn Gly Glu Trp Tyr Leu Ile Gly Met Pro Gly Lys Lys Asn Thr Ile
195 200 205

Thr Tyr Ala Cys Cys Pro Glu Pro Tyr Val Asp Val Thr Phe Thr Ile
210 215 220

Met Ile Arg Arg Arg Thr Leu Tyr Tyr Phe Phe Asn Leu Ile Val Pro
225 230 235 240

Cys Val Leu Ile Ser Ser Met Ala Leu Leu Gly Phe Thr Leu Pro Pro
245 250 255

Asp Ser Gly Glu Lys Leu Thr Leu Gly Val Thr Ile Leu Leu Ser Leu
260 265 270

Thr Val Phe Leu Asn Leu Val Ala Glu Thr Leu Pro Gln Val Ser Asp
275 280 285

Ala Ile Pro Leu Leu Gly Thr Tyr Phe Asn Cys Ile Met Phe Met Val
290 295 300

Ala Ser Ser Val Val Leu Thr Val Val Val Leu Asn Tyr His His Arg
305 310 315 320

Thr Ala Asp Ile His Glu Met Pro Gln Trp Ile Lys Ser Val Phe Leu
325 330 335

Gln Trp Leu Pro Trp Ile Leu Arg Met Ser Arg Pro Gly Lys Lys Ile
340 345 350

Thr Arg Lys Thr Ile Met Met Asn Thr Arg Met Arg Glu Leu Glu Leu
355 360 365

Lys Glu Arg Ser Ser Lys Ser Leu Leu Ala Asn Val Leu Asp Ile Asp
370 375 380

Asp Asp Phe Arg His Gly Pro Pro Pro Pro Asn Ser Thr Ala Ser Thr
385 390 395 400

Gly Asn Leu Gly Pro Gly Cys Ser Ile Phe Arg Thr Asp Phe Arg Arg
405 410 415

Ser Phe Val Arg Pro Ser Thr Met Glu Asp Val Gly Gly Gly Leu Gly
420 425 430

Ser His His Arg Glu Leu His Leu Ile Leu Arg Glu Leu Gln Phe Ile
435 440 445

Thr Ala Arg Met Lys Lys Ala Asp Glu Glu Ala Glu Leu Ile Ser Asp
450 455 460

Trp Lys Phe Ala Ala Met Val Val Asp Arg Phe Cys Leu Phe Val Phe
465 470 475 480

Thr Leu Phe Thr Ile Ile Ala Thr Val Ala Val Leu Leu Ser Ala Pro
485 490 495

His Ile Ile Val Gln
500

Patent Claims

1. Nucleic acid which comprises a sequence selected from

(a) the sequences according to SEQ ID NO: 1, SEQ ID NO: 3 or SEQ ID NO: 5,

(b) part sequences, which are least 14 base pairs in length, of the sequences defined under (a),

(c) sequences which hybridize with the sequences defined under (a) in 2 x SSC at 60°C, preferably in 0.5 x SSC at 60°C, particularly preferably in 0.2 x SSC at 60°C,

(d) sequences which exhibit at least 70% identity with the sequences defined under (a), between position 1295 and position 2195 from SEQ ID NO: 1, or between position 432 and position 1318 from SEQ ID NO: 3, or between position 154 and position 1123 from SEQ ID NO: 5,

(e) sequences which are complementary to the sequences defined under (a), and

(f) sequences which, on account of the degeneracy of the genetic code, encode the same amino acid sequences as the sequences defined under (a) to (d).

2. Vector which comprises at least one nucleic acid according to Claim 1.

3. Vector according to Claim 2, characterized in that the nucleic acid molecule is functionally linked to regulatory sequences which ensure the expression of the nucleic acid in prokaryotic or eukaryotic cells.

4. Host cell which contains a nucleic acid according to Claim 1 or a vector according to Claim 2 or 3.

5. Host cell according to Claim 4, characterized in that it is a prokaryotic or eukaryotic cell.

6. Host cell according to Claim 5, characterized in that the prokaryotic cell is E.coli.

7. Host cell according to Claim 5, characterized in that the eukaryotic cell is a mammalian cell or an insect cell.

8. Polypeptide which is encoded by a nucleic acid according to Claim 1.

9. Acetylcholine receptor which comprises at least one polypeptide according to Claim 8.

10. Process for preparing a polypeptide according to Claim 8, which comprises

(a) culturing a host cell according to one of Claims 4 to 7 under conditions which ensure the expression of the nucleic acid according to Claim 1, and

(b) isolating the polypeptide from the cell or the culture medium.

11. Antibody which reacts specifically with the polypeptide according to Claim 8 or the receptor according to Claim 9.

12. Transgenic invertebrate which contains a nucleic acid according to Claim 1.

13. Transgenic invertebrate according to Claim 12, characterized in that it is Drosophila melanogaster or Caenorhabditis elegans.

14. Process for producing a transgenic invertebrate according to Claim 12 or 13, which comprises introducing a nucleic acid according to Claim 1 or a vector according to Claim 2 or 3.

15. Transgenic progeny of an invertebrate according to Claim 12 or 13.

16. Process for preparing a nucleic acid according to Claim 1, which comprises the following steps:

(a) carrying out an entirely chemical synthesis in a manner known per se, or

(b) chemically synthesizing oligonucleotides, labelling the oligonucleotides, hybridizing the oligonucleotides to the DNA of an insect cDNA library, selecting positive clones and isolating the hybridizing DNA from positive clones, or

(c) chemically synthesizing oligonucleotides and amplifying the target DNA by means of PCR.

17. Regulatory region which naturally controls transcription of a nucleic acid according to Claim 1 in insect cells and ensures specific expression.

18. Process for discovering novel active compounds for plant protection, in particular compounds which alter the conducting properties of receptors according to Claim 9, which comprises the following steps:

(a) providing a host cell according to one of Claims 4 to 7,

(b) culturing the host cell in the presence of a compound or a sample which comprises a multiplicity of compounds, and

(c) detecting altered performance properties.

19. Process for discovering a compound which binds to receptors according to Claim 9, which encompasses the following steps:

(a) bringing a host cell according to one of Claims 4 to 7, a polypeptide according to Claim 8 or a receptor according to Claim 9 into contact with a compound or a mixture of compounds under conditions which permit interaction of the compound(s) with the host cell, the polypeptide or the receptor, and

(b) determining the compound(s) which bind(s) specifically to the receptors.

20. Process for discovering compounds which alter the expression of receptors according to Claim 9, which comprises the following steps:

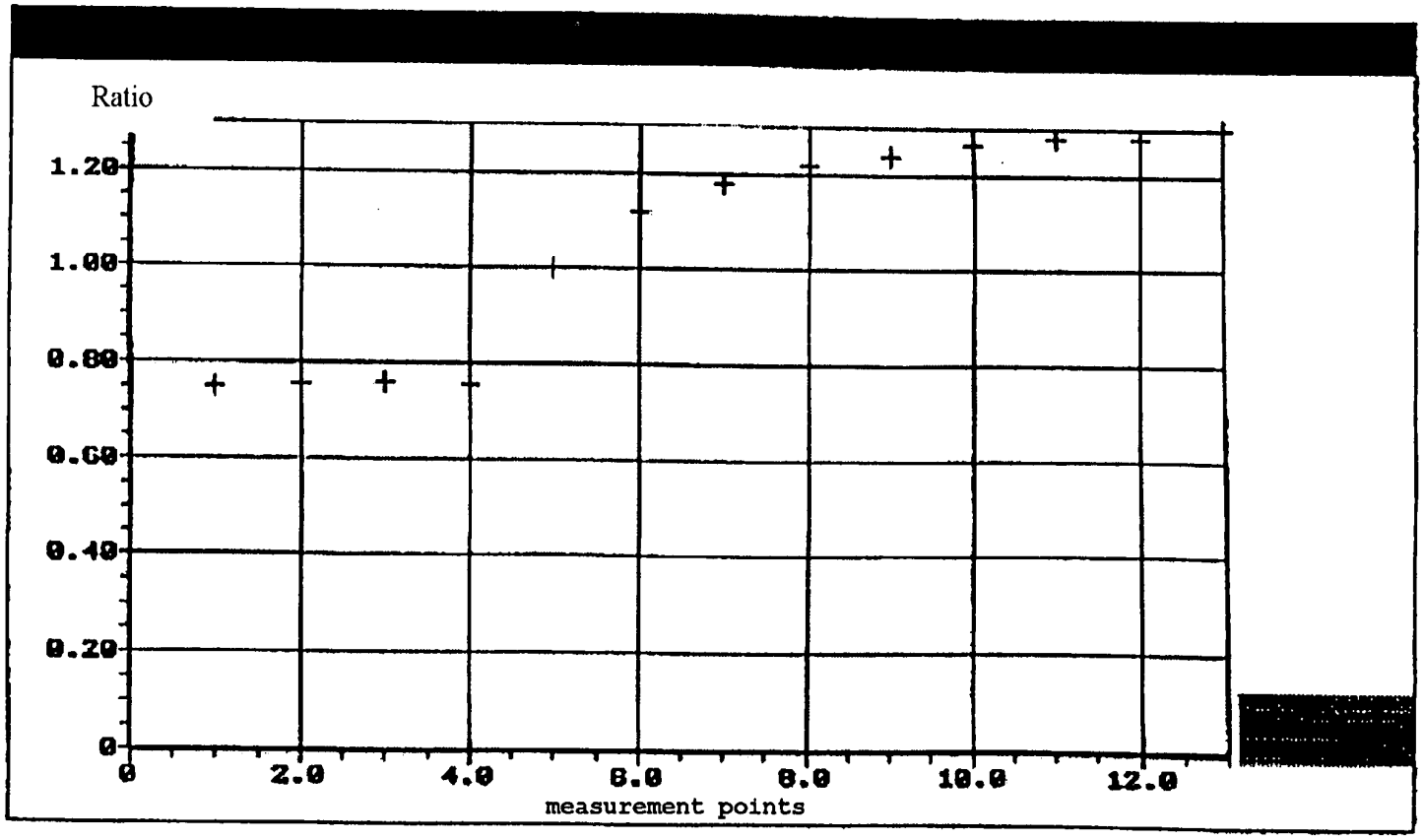
(a) bringing a host cell according to one of Claims 4 to 7 or a transgenic invertebrate according to Claim 11 or 12 into contact with a compound or a mixture of compounds,

(b) determining the receptor concentration, and

(c) determining the compound(s) which specifically influence(s) the expression of the receptor.

21. Use of at least one nucleic acid according to Claim 1, one vector according to Claim 2 or 3, one regulatory region according to Claim 16 or one antibody according to Claim 11 for discovering novel active compounds for plant protection or for discovering genes which encode polypeptides which are involved in synthesizing functionally similar acetylcholine receptors in insects.

Figure 1


Ratio
1.20
1.00
0.80
0.60
0.40
0.20
0
0
2.0
4.0
6.0
8.0
10.0
12.0
measurement points